AMENDMENT

Please cancel claims 1-45 without prejudice as to their further prosecution, and add new claims 46 - 92 as follows:

46. (New) An automated method for identifying a component in a biological sample, comprising:

using a mass spectrometer to generate a computer readable data set comprising data representing components in the biological sample for analysis by a computer, and using the computer to:

denoise the data set to generate denoised data;

correct a baseline from the denoised data to generate an intermediate data set;

define putative peaks in the intermediate data set, wherein the putative peaks represent components in the biological sample;

generate a residual baseline by removing the putative peaks from the intermediate data set;

remove the residual baseline from the intermediate data set to generate a corrected data set;

locate a probable peak in the corrected data set; and

identify the component that corresponds to the located probable peak.

47. (New) The method of claim 46, wherein said mass spectrometer is a MALDI-TOF mass spectrometer.

48. (New) The method according to claim 46 wherein denoising the data set includes generating a noise profile for the data set.

49. (New) The method according to claim 46 wherein denoising the data set includes transforming the data set using wavelet technology into a series of stages.

50. (New) The method according to claim 49 further including generating a noise profile for stage 0.

51. (New) The method according to claim 50 further including generating a noise profile for other stages.

52. (New) The method according to claim 51 wherein the noise profile for each of the other stages stage 0 scaled by a scaling factor

53. (New) The method according to claim 52 wherein the scaling factor is derived from the end portion of each of the other stages, respectively.

54. (New) The method according to claim 49 further including applying a threshold to selected stages, the threshold being derived from the noise profile.

55. (New) The method according to claim of 54 wherein the threshold is scaled by a threshold factor before being applied to the selected stages.

56. (New) The method according to claim 55 wherein the threshold factor is selected so that higher stages of data are filtered less than lower stages.

57. (New) The method according to claim 49 further including generating a sparse data set indicative of the denoised data.

58. (New) The method according to claim 49 further including shifting the denoised data to account for variations due to a starting value for the wavelet transformation.

59. (New) The method according to claim 46 wherein correcting the baseline further includes generating a moving average of the denoised data set.

60. (New) The method according to claim 59 wherein the moving average is used to find peak sections in the denoised data set.

61. (New) The method according to claim 60 wherein the peak sections are removed from the denoised data set.

62. (New) The method according to claim 61 further including generating a baseline correction.

63. (New) The method according to claim 46 further including compressing the intermediate data set, the intermediate data set having a plurality of data values associated with respective addresses.

64. (New) The method according to claim 63 wherein a compressed data value is a real number that includes a whole portion representing the difference between two addresses.

65. (New) The method according to claim 63 wherein a compressed data value is a real number that includes a decimal portion representing the difference between a maximum value of all the data values and a value at a particular address.

66. (New) The method according to claim 46 further including performing a mass shift based on the position of the putative peaks.

67. (New) The method according to claim 46 wherein generating the residual baseline includes deleting an area around each putative peak in the intermediate data set.

68. (New) The method according to claim 67 wherein the area deleted is derived from a determined width of a putative peak.

69 (New) The method according to claim 67 wherein the residual baseline is derived from data remaining in the intermediate data set after the areas around the putative peaks have been removed.

70. (New) The method according to claim 69, wherein an area equal to twice the width of the Gaussian is removed from the left of the center line of the putative peaks.

71. (New) The method according to claim 69, wherein an area equivalent to 50 daltons is removed from the right of the center line of the putative peaks.

72. (New) The method according to claim 67 wherein generating the residual baseline includes fitting a quartic polynomial to the data set remaining in the intermediate data after the peaks have been removed.

73. (New) The method according to claim 46 wherein the probable peak is located by fitting a Gaussian curve to a peak area in the corrected data set.

74. (New) The method according to claim 46 wherein the identifying step includes using a generated noise profile to calculate the signal-to-noise ratio for the probable peak.

75. (New) The method according to claim 74 wherein a residual peak error is calculated by comparing the probable peak to a Gaussian curve.

76. (New) The method according to claim 75 wherein the residual peak error is used to adjust the signal-to-noise ratio to generate an adjusted signal-to-noise ratio.

77. (New) The method according to claim 46 wherein the identifying step includes

deriving a peak probability for the probable peak.

78. (New) The method according to claim 77 wherein the peak probability is derived using the signal-to-noise ratio.

79. (New) The method according to claim 78 wherein the peak probability is derived by using an allelic ratio, the allelic ratio being a comparison of two peak heights indicated in the corrected data.

80. (New) The method according to claim 46 wherein the identifying step includes calculating a peak probability that a probable peak in the corrected data is a peak indicating composition of the biological sample.

81. (New) The method according to claim 80 wherein peak probability is calculated for each of a plurality of probable peaks in the corrected data.

82. (New) The method according to claim 81 wherein a highest probability is compared to a second-highest probability to generate a calling ratio.

83. (New) The method according to claim 82 wherein the calling ratio is used to determine if the composition of the biological sample will be called.

84. (New) A computerized system for identifying a component in a biological sample, the system comprising:

an instrument for receiving the biological sample and generating a data set capable of analysis by a computer comprising data representing components in the biological sample;

a computer communicating with the instrument and configured to receive the generated data set, the computer programmed to perform the method of:

denoising the data set to generate denoised data;

correcting a baseline from the denoised data to generate an intermediate data set;

defining putative peaks in the intermediate data set, wherein the putative peaks represent components in the biological sample;

generating a residual baseline by removing the putative peaks from the intermediate data set;

removing the residual baseline from the intermediate data set to generate a corrected data set;

locating a probable peak in the corrected data set; and

identifying the component that corresponds to the located probable peak.

85. (New) The system according to claim 84 wherein the computer is integral to the instrument.

86. (New) A machine readable program operating on a computing device, the computing device being configured to receive a data set comprising computer readable data representing components of a biological sample, wherein the program directs the computing device to implements the steps of:

denoising the data set to generate denoised data;

correcting a baseline from the denoised data to generate an intermediate data set;

defining putative peaks ~~for~~ in the intermediate data set, wherein the putative peaks represent components in the biological sample;

generating a residual baseline by removing the putative peaks from the intermediate data set;

removing the residual baseline from the intermediate data set to generate a corrected data set;

locating a probable peak in the corrected data set; and

identifying the component that corresponds to the located probable peak.

87. (New) A system for identifying a component of a DNA sample, comprising:

a mass spectrometer for receiving the DNA sample and generating a computer readable data set comprising data representing components in the DNA sample;

a computing device configured to receive the computer readable data set, the computing device programmed to implement~~ing~~ the method comprising:

denoising the data set to generate denoised data;

correcting ~~the~~ a baseline from the denoised data to generate an intermediate data set;

defining ~~the~~ putative peaks ~~for~~ in the intermediate data set, wherein the putative peaks represent components in the biological sample;

generating a residual baseline by removing the putative peaks from the intermediate data set;

removing the residual baseline from the intermediate data set to generate a corrected data set;

locating a probable peak in the corrected data set; and

using the located probable peak to identify the component that corresponds thereto.

88. (New) The system according to claim 87, where the method further includes using a statistical methodology to determine if the located probable peak is an actual peak.

89. (New) The system according to claim 88, where the method further includes determining whether a probability of the actual peak existing is sufficiently high to identify the component of the DNA sample, and if the probability is not sufficiently high, then the method does not identify the component.

90. (New) The system according to claim 89, where the percentage of correctly called components is about 100 percent.

91. (New) A system for identifying a component in a biological sample, comprising:

an instrument receiving the biological sample and generating a computer readable data set indicative of the component in the biological sample;

a computing device for receiving the computer readable data set and performing the steps of:

generating a corrected data set by processing the data set to remove noise due to system and chemical reaction characteristics, the corrected data set having putative peak areas;

defining the position of expected peaks using known possible peak areas from the biological sample;

shifting the corrected data set to more closely align the putative peaks to the expected peaks;

calculating the probability that each of the putative peaks in the shifted data set are actual peaks;

comparing the highest probability to the second-highest probability to generate a calling ratio; and

using the calling ratio to determine if the identity of the component of the biological sample is called.